Received: 16 August2011 ■ *Revised: 225 November 2011* ■ *Accepted: 1 December 2011* ■ *Publication: 3 January 2012*

JNIRS JOURNAL OF NEAR INFRARED SPECTROSCOPY

Developing robust faecal near infrared spectroscopy calibrations to predict diet dry matter digestibility in cattle consuming tropical forages

D.B. Coates[a] and R.M. Dixon[b]

[a]CSIRO Ecosystem Sciences, ATSIP, PMB, PO, Aitkenvale, Townsville, Queensland 4814, Australia. E-mail: coatesdc@bigpond.net.au

[b]The University of Queensland, QAAFI, PO Box 6014, Rockhampton, Queensland 4701, Australia

Various studies, mainly from temperate areas, have reported calibrations developed from the near infrared (NIR) spectra of faeces (F.NIRS) for predicting diet digestibility in ruminants and there has been substantial variation in predictive accuracy as indicated by calibration and validation statistics. The present study was conducted to develop and examine the reliability and robustness of F.NIRS calibration equations to estimate dry matter digestibility (DMD) of forage diets ingested by cattle grazing in the rangelands of northern Australia. A large and diverse calibration data set of matched diet–faecal pairs was obtained over 10 years using three sampling methods: (1) grazed pasture with diet samples collected from oesophageal fistulated steers and faeces collected from resident cattle; (2) *in vivo* digestibility experiments with penned cattle fed forage hays; and (3) penned cattle fed pasture freshly harvested from the field. Estimated *in vivo* DMD reference values were determined using pepsin–cellulase *in vitro* analysis of diet samples. The final calibration set of 1052 samples represented 264 diets with DMD ranging from 38% to 75%. Calibration statistics for DMD% were: standard error of calibration = 1.87, standard error of cross validation = 1.91, and the coefficient of determination, r^2 = 0.90. Factors of particular importance, with regard to the accuracy of DMD reference values, are identified and discussed and recommendations made for minimising reference errors. A comprehensive series of independent validation tests was conducted by selecting validation sample sets from the entire sample set according to a range of criteria. Each validation sub-set was tested using the calibration calculated from the remainder of the sample set. These tests showed that sampling method and experimental site often had important effects on calibration statistics and performance and also that the standard error of performance of the overall calibration would likely be <2.5 DMD percentage units when applied to samples sourced from regions and pasture types represented in the calibration. Despite the large size and diversity of the calibration data set it was concluded that robustness would likely be improved by expansion of the calibration data set.

Keywords: faecal NIR spectroscopy, grazing cattle, digestibility, robustness, miss-match error, validation, northern Australia

Introduction

The energy value of forages as feedstuffs for ruminants depends primarily on the digestibility of dry matter or organic matter, where digestibility (apparent rather than true digestibility) is defined as the proportional difference between the amount consumed and the amount excreted in the faeces.[1] Thus, there has been enormous effort by animal nutritionists to develop methods to accurately measure digestibility. The most comprehensive and direct approach involves measure-

ISSN: 0967-0335
doi: 10.1255/jnirs.967

ment of the amount of feed ingested and faeces excreted by total collection but, as well as being laborious and costly, it can only be applied to discrete feeds offered to animals in pens. Therefore, alternative techniques have been developed to estimate digestibility using laboratory procedures. The most common *in vitro* method involves the incubation of ground feeds with either rumen fluid[2] or with pepsin and cellulase solutions.[3,4] A variation is to incubate small synthetic-fibre bags containing forage in the rumen of animals surgically prepared with rumen canulae.[5,6] Other laboratory methods attempted to use regression relationships between fractions or chemical components of the feedstuff (for example, lignin, neutral detergent fibre, acid detergent fibre) and digestibility.[1,7–9] However, these regression relationships are not sufficiently robust for general application.

Measuring the digestibility of the diet selected by grazing animals presents additional difficulties because the diet selected usually differs substantially from that of the pasture on offer. In most situations, it is not possible to manually harvest plant material that accurately represents the diet selected by the grazing animal. This led to the development and adoption of sampling procedures using oesophageal fistulated (OF) animals so that samples of selectively grazed forage (extrusa) representing the diet of the grazing animals could be collected and analysed in the laboratory.[10,11] However, this approach may also involve major errors since samples of extrusa often do not accurately represent the diet selected by the resident grazing animals.[12–14]

The development of near infrared (NIR) reflectance spectroscopy provided new scope for making low cost and accurate estimates of digestibility of feedstuffs, including forages.[15–18] In addition, the development of faecal NIR (F.NIRS) technology, where NIR spectra of faeces are used to predict diet attributes, allows estimation of the digestibility of diets selected by grazing animals.[19–23]

In 1994, we commenced a programme to develop a F.NIRS calibration equation to predict dry matter digestibility (DMD) in cattle grazing tropical and sub-tropical pasture systems in northern Australia. We identified two major challenges. One was to ensure the accuracy of estimated *in vivo* digestibility reference values using *in vitro* analysis. The second was to achieve robustness by constructing a calibration data set which would encompass the range of environments and pasture systems found in the northern half of the Australian continent. Three sampling methods were used to build a calibration sample set that was appropriate for grazing cattle and provided the desired diversity of samples in relation to the pasture, diet, soil, climatic, seasonal and year effects likely to be encountered in the most important pasture systems in the northern Australian rangelands. This paper describes the methods used to generate appropriate and diverse calibration samples, procedures adopted to ensure the accuracy of reference values, the derived calibration equations and an assessment of calibration robustness based on comprehensive validation tests.

Materials and methods

Sampling methods

Developing F.NIRS calibrations for predicting diet attributes such as digestibility requires samples of the diet from which reference values are determined by laboratory analysis and faeces from animals consuming the relevant diet for measurement of the NIR spectra (i.e. diet–faecal pairs). We used three sampling methods to generate the required diet–faecal pairs: (1) sampling from grazed pastures (GP), (2) conventional *in vivo* digestibility pen trials (VIVO) and (3) fresh feed short-term pen trials (FR-FEED). Calibration samples were generated over the ten year period 1994 to 2003.

Grazed pastures (GP sample set)

Initially (1994–1997), most diet–faecal pairs were generated from grazed pastures. Forage samples to represent the diet of the resident cattle were obtained by collecting extrusa from three to six OF steers grazing each paddock at each sampling event. The OF steers at each location were familiar with the pastures being grazed and only one sample per sampling event was collected from each OF steer. Faecal samples were obtained from the three to five resident animals by rectal sampling in yards or from fresh faecal pats in the paddock. Although it would have been preferable to collect the faecal samples one to two days after sampling with the OF steers, due to various constraints the OF extrusa and faecal samples were collected on the same day. However, paddock size and pasture dry matter on offer were sufficient to ensure that there was little change in the quality of the diet selected over a two day interval. To enhance similarity between extrusa collected from the OF steers and the diet selected by resident cattle, sampling was restricted to paddocks of no more than 5 ha with uniform pasture composition and a limited number of component species.

Diversity in the calibration set with respect to pasture species, climate and soils was achieved by sampling a range of pastures at five different Queensland locations (see Table 1): (1) Lansdown Research Station (LDN) south of Townsville with sown grass/legume pastures growing on coastal solodic soils; (2) Cardigan Station (CDG) south of Charters Towers with native and sown grass/legume pastures growing on red duplex soils; (3) Hillgrove Station (HLG) north of Charters Towers with native and sown grass/legume pastures growing on black basaltic soil; (4) Springmount Station (SPR) west of Mareeba with native and grass/legume pastures growing on low fertility red duplex soil derived from granite; and (5) Brian Pastures Research Station (BP) near Gayndah with native pastures and a range of sown species growing on fertile brown cracking clays. Sampling was conducted at various times through the year and for a number of years to encompass the range in pasture quality associated with the annual seasonal cycle and between-year differences. In total, the sampling of grazed pastures (GP sample set) represented 116 forage diets and, with animal replication, contributed 468 faecal spectra to the calibration set, as summarised in Table 2.

Table 1. Experimental sites used for generating diet-faecal pairs from grazed pasture experiments and fresh feed pen trials showing site codes and site co-ordinates.

Site	Site code	Site co-ordinates
Lansdown Research Station	LDN	19° 41′ S, 146° 51′ E
Cardigan Station	CDG	20° 12′ S, 146° 42′ E
Hillgrove Station	HLG	19° 38′ S, 145° 48′ E
Springmount Station	SPR	17° 13′ S, 145° 18′ E
Brian Pastures	BP	25° 40′ S, 151° 45′ E
Swans Lagoon	SWL	19° 41′ S, 146° 51′ E
Brunchilly Station	BCH	18° 52′ S, 134° 30′ E
Katherine Research Station	KTH	14° 28′ S, 132° 18′ E

Conventional *in vivo* digestibility trials (VIVO sample set)

Forage hays selected for batch uniformity were chaffed and fed to individually penned cattle for a nine day adaptation period followed by an eight day total faecal collection period. Most of the hays were purchased from commercial sources and most were introduced pasture grasses or legumes. Each diet was fed to between two and ten young cattle but usually to three or four. Although actual apparent *in vivo* digestibility was measured in these trials, the DMD reference values used for this study were determined by *in vitro* analysis of sub-samples taken from the forage offered during the eight day faecal collection period and bulked over days within diet. Forage refusals for the final eight days were collected from individual animals for subsequent analysis to enable DMD reference values to be adjusted to the diet consumed. Faecal spectra were derived from faecal samples bulked over the final three days of each trial for each animal. These trials contributed 71 diets and, with the animal replication, 296 faecal spectra to the calibration set. The diets included hays of 19 grass species (four native C_4 grasses, 12 introduced C_4 grasses and three introduced C_3 grasses), seven pasture legumes, five grass/legume mixtures and two grass species mixtures.

Fresh feed short-term pen trials (FR-FEED sample set)

Trials were conducted with individually penned cattle fed forage harvested direct from the paddock using a tractor-drawn forage harvester or other mechanical means or, for a few diets, by hand harvesting. Green forage was harvested shortly before feeding, either once daily or twice daily, the latter protocol being adopted when pasture was at an early stage of growth with high moisture content and when the chemical composition was likely to change rapidly after harvest. Mature, dry pasture was harvested less frequently and stored under cover. Areas of pasture to be harvested were carefully selected for uniformity of species composition and stage of growth to minimise differences between meals in diet quality. Only data from trials where forage could be fed for the desired interval without interruption by rain were used.

Table 2. Contribution to the calibration data set of diet–faecal pairs from grazed pasture according to pasture types, sites, and the number of different diets.

Pasture type	Locations[a]	Diets
Native grass pasture	SPR, CDG, HLG, BP	18
Native pasture/stylo (Stylosanthes hamata and S. scabra)	LDN, SPR	22
Urochloa (Urochloa mosambicensis)/stylo	LDN, CDG	44
Buffel (Cenchrus ciliaris)/stylo	LDN, CDG	16
Buffel	HLG	2
Verano stylo (S. hamata cv. Verano)	LDN	5
Seca stylo (S. scabra cv. Seca)	LDN	6
Butterfly pea (Clitoria ternatea)	BP	1
Lab lab (Lab lab purpureus)	BP	1
Creeping blue grass (Bothriochloa insculpta)	BP	1

[a] See Table 1 for the site codes.

To achieve matched diet–faecal pairs. it was necessary to feed diets for sufficient time to allow faecal composition to equilibrate with the test diet. Preliminary trials indicated that a feeding duration of five days was sufficient, although most of the trials were of six to ten days duration. Each diet was fed to three to five young cattle. Representative samples of the diet were obtained at each meal and reference values were calculated as the mean of forage samples collected over the three days prior to the collection of calibration faecal samples. Faecal samples were obtained by sampling fresh faeces from the floor of pens and calibration faecal samples were restricted to those collected after faeces had equilibrated with the diet, usually those from the final one to three days of each trial. Equilibration was assessed by plotting predicted dietary nitrogen concentration and digestibility against day using earlier versions of the F.NIRS calibration equations.

FR-FEED trials were conducted primarily at four sites, two in the northern speargrass pasture region of NE Queensland (LDN and Swans Lagoon Research Station (SWL) 100 km SSE of Townsville), and two sites in the Northern Territory (Brunchilly Station (BCH) on the Barkly Tableland and Katherine Research Station (KTH) near Katherine) (also see Table 1). This was intended to provide diversity in pasture species and species mixtures in the diets as well as in climate, soil and seasonal conditions. FR-FEED trials contributed 77 diets and 288 faecal spectra to the calibration set.

Reference values

In vitro dry matter loss (IVDML) was determined, in duplicate, on samples of the forages offered (VIVO and FR-FEED trials) or extrusa samples (GP sample set) using a modification[24] of the two-stage pepsin–cellulase procedure of Jones and Hayward.[3] Values for extrusa samples were adjusted to correct for (1) the effect of saliva on the pepsin–cellulase digestion of forage[24] and (2) for possible miss-match error.[25] We define miss-match error as the error in reference values caused by differences between the composition of samples analysed to derive reference values (i.e. samples of extrusa or forage offered in pens) and the composition of the diet of the cattle from which the faecal spectra are derived. This dual adjustment was considered essential for providing reliable reference IVDML values for all grazed pasture diets and has been described comprehensively by Coates.[25] Finally, estimated *in vivo* DMD (hereafter designated as DMD) was calculated from adjusted IVDML values, also according to the method described by Coates.[25]

Faecal near infrared spectroscopy analysis, calibrations and validations

Faecal samples were ovendried at 65°C and then ground through a 1 mm screen (Model 1093 Cyclotec Mill, Foss Tecator AB, Hoganas, Sweden). Samples were subsequently re-dried at 65°C overnight, cooled in a desiccator, placed in a NIR spinning cup module, and scanned in the 400–2500 nm range using a scanning monochromator (Model 6500; NIRSystems, Inc., Silver Spring, MD, USA). Calibration equations were developed using ISI software (Infrasoft International, Port Matilda, PA, USA). Standard normal variate and detrend transformations were applied to the first derivative of the Log(1/*R*) absorbance data (1,4,4,1) over the 700–2500 nm wavelength band and regression equations were calculated using modified partial least squares. Critical *"t"* and *"H"* outlier values were set at 3 and 9, respectively, and where the critical values were exceeded, the samples were eliminated from the calibration data as outliers. Three outlier elimination passes were used.

An overall calibration equation, ALL.EQA, $n = 1052$, was developed using the data sets from all three sampling methods. The complete data set comprised 264 diets and, on average, there were four animal replicate faecal spectra per diet. Reference values for replicate spectra were generally the same within the VIVO and FR-FEED data sets because between animal differences within diet were usually negligible after adjusting DMD on the basis of individual refusals. However, reference values often differed between replicate faecal spectra for pastures sampled in the GP data set due to cattle selecting diets that differed in grass/non-grass proportions.[24] Replicate spectra were used to prevent any animal effect confounding the resultant calibration equation. Equations were also developed for each of the GP, VIVO and FR-FEED sample sets (GP.EQA, VIVO.EQA and FR-FEED.EQA, respectively) to investigate any special strengths or weaknesses associated with sampling method. In addition, the relationships between reference and predicted DMD for different groupings of samples (for example, according to diet type or geographical location) were also investigated to detect any specific problems regarding predictive accuracy or bias in relation to such groupings.

Predictive accuracy and robustness were investigated by dividing the full calibration sample set into sub-sets according to (1) sampling method and (2) sites within sampling method. Validation tests were then performed where the validation sample set was independent of the respective calibration set. Outliers eliminated ($n = 8$) during the main calibration (ALL) were not included in the validation sets so that validation statistics could be compared on an equal basis with calibration statistics. Thus, calibration equations developed on the GP, VIVO and FR-FEED sample sets were used to predict digestibility on the other sample sets and the validation statistics computed. Similarly, another set of calibrations was developed by excluding GP or VIVO or FR-FEED samples, in turn, from the entire sample set to create equations ALLlessGP.EQA, ALLlessVIVO.EQA and ALLlessFF.EQA. These equations were also used to predict DMD of samples in the GP, VIVO and FR-FEED sets, respectively, and validation statistics computed.

This validation procedure was extended further by developing calibration equations on sample sub-sets from the GP sample set after excluding samples from either Cardigan (GPlessCDG.EQA), Springmount (GPlessSPR.EQA), Brian Pastures (GPlessBP.EQA), or Lansdown (GPlessLDN.EQA) and validation statistics were derived using GPlessCDG.EQA to predict the Cardigan samples and so forth. In a similar manner, the process was repeated by excluding Cardigan,

Table 3. Faecal near infrared spectroscopy calibration statistics for the prediction of estimated in vivo dry matter digestibility (%) according to the sampling procedure used to generate diet–faecal pairs.

Sample set[a]	*n*[b]	Outliers[c]	Range	*SEC*	*SECV*	R^2
ALL	1052	8	38–75	1.87	1.91	0.90
GP set	468	5	47–75	1.53	1.61	0.91
VIVO set	296	5	44–75	1.46	1.63	0.95
FR-FEED set	288	1	38–60	1.09	1.22	0.90
ALLlessGP	584	3	38–75	1.56	1.64	0.92
ALLlessVIVO	756	7	38–75	1.73	1.79	0.90
ALLlessFF	764	4	44–75	1.91	1.97	0.90

[a] See text for sample set codes.
[b] Number of samples in calibration set including outliers not contributing to the equation.
[c] Outliers eliminated during calibration and not contributing to the calibration equation.
SEC, standard error of calibration; *SECV*, standard error of cross validation.

Springmount, Brian Pastures, or Lansdown samples in turn from the entire sample set to produce calibration equations ALLlessCDG.EQA, ALLlessSPR.EQA, ALLlessBP.EQA and ALLlessLDN.EQA and then repeating the validation exercises on the Cardigan, Springmount, Brian Pastures and Lansdown samples. The same procedure was applied to the FR-FEED (shown as FF) sample set with calibration equations FFlessKTH.EQA, ALLlessKTH.EQA, FFlessSWL.EQA, ALLlessSWL.EQA, FFlessBCH.EQA, ALLlessBCH.EQA, FFlessLDN.EQA and ALLlessLDN.EQA being used to perform validation tests on samples from Katherine, Swans Lagoon, Brunchilly and Lansdown.

The purpose of the above validation tests was to gain an understanding of the influence of variables such as sampling method, site and climatic factors on calibration robustness rather than to provide a validation test on the final calibration equation ALL.EQA. This final test was conducted by dividing the full sample set into three similar sized groups, Grp-A, Grp-B and Grp-C. Each of these groups contained approximately the same number of samples from each sampling method and each GP or FR-FEED site and each group represented similar ranges with regard to seasons and diet DMD reference values. Calibration sub-sets were developed by combining samples from two groups to give equations GrpA + GrpB.EQA, GrpA + GrpC.EQA and GrpB + GrpC.EQA and these equations were then used for validation tests on GrpC, GrpB and GrpA respectively. Faecal samples from any given diet (i.e. animal replications) were allocated to the same group to ensure that Grp-A, Grp-B and Grp-C validation sets were independent of the calibration sets.

To gain some insight into the potential improvement of predictive accuracy that might be achieved by increasing the number and diversity of diets represented in the calibration set, prediction statistics (computed on the same basis as validation statistics) were also derived for all the above sets and sub-sets (GP, VIVO, FR-FEED, Cardigan(GP), Springmount(GP), Brian Pastures(GP), Lansdown(GP), Katherine(FF), Swans Lagoon(FF), Brunchilly(FF) and Lansdown(FF), Grp A, Grp B and Grp C) for predictions made using the overall calibration ALL.EQA. Clearly such measures of predictive accuracy did not constitute independent validation but any improvement in predictive accuracy afforded by ALL.EQA, compared with the performance of the various restricted calibration equations, likely indicated potential for an expanded calibration set to improve robustness through the inclusion of more and different diets.

Regression standard errors of calibration (*SEC*s) and standard errors of performance (*SEP*s) were used as the main measures of predictive accuracy. The coefficient of determination (r^2) was deemed to be of limited value because of the large effect of the attribute range on r^2 values and the large differences between sample sub-sets in the range of DMD reference values.

Animal welfare

All experiments were conducted with the approval of the relevant Animal Ethics Committee operating at the time the different experiments were performed.

Results and discussion

Reference digestibilities and distribution

The range in DMD reference values differed between sampling methods with ranges of 31, 28 and 22 percentage units for the VIVO, GP and FR-FEED sets, respectively (Table 3). Frequency distributions of reference values for the three sampling methods are shown in Figure 1. Of particular concern was the narrow range and generally low DMD values in the FR-FEED set with few values > 55% and all < 60%. DMD values ⩾ 60% would have been absent from the VIVO sample set as well if temperate grasses and legumes had not been among the hays fed. Thus, the only sampling method that provided high DMD values for tropical forages was the GP method where 26.5% of calibration samples were from diets with DMD > 60%. For the entire sample set, 22% had DMD values < 50%, 61%

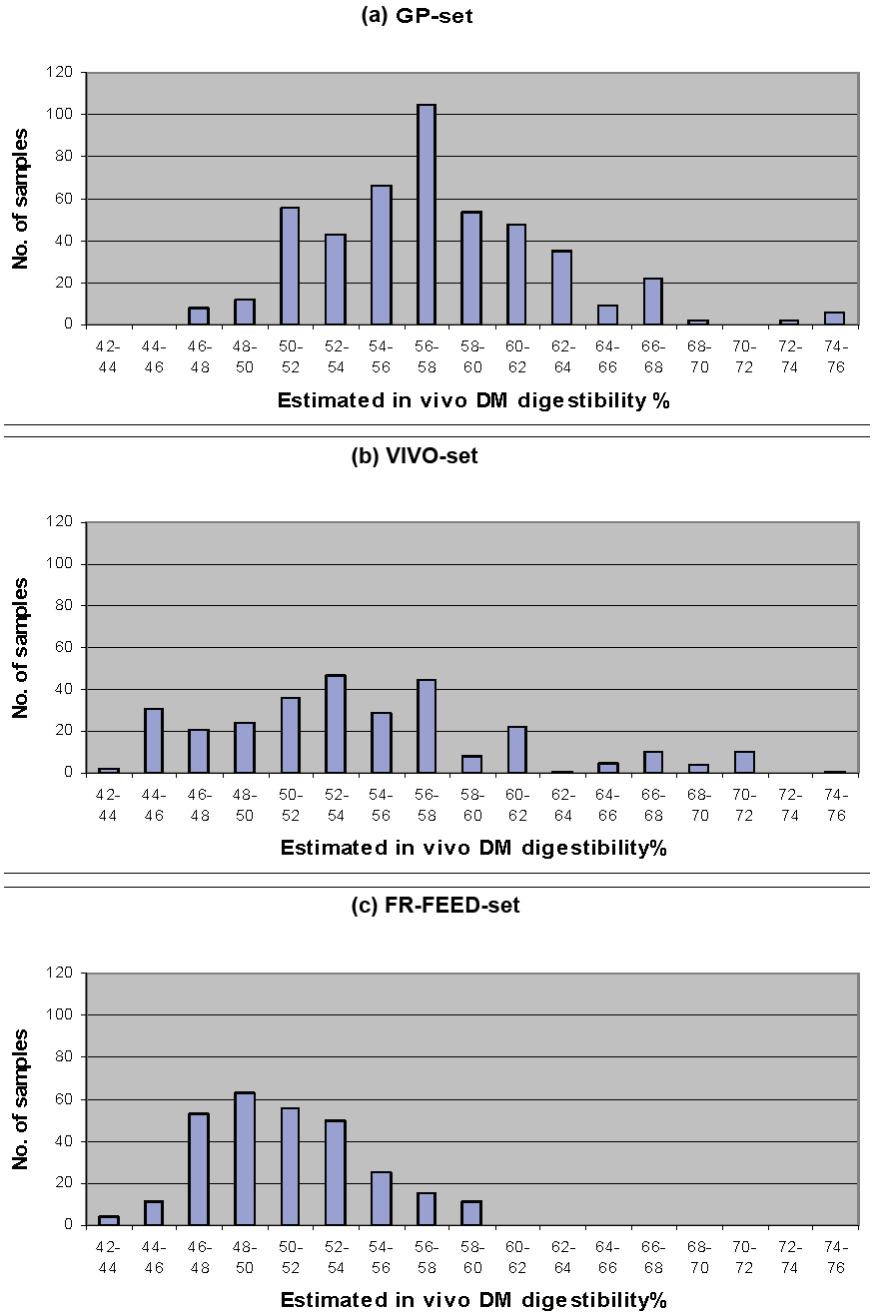

Figure 1. Frequency distribution of samples according to DMD reference values and sampling method. (a) GP sample set, (b) VIVO sample set, (c) FR-FEED sample set.

had DMD in the 50–60% range and 17% had DMD > 60%. This frequency distribution in the full sample set was considered appropriate for cattle grazing pastures in northern Australia but it was clear that there would have been a serious lack of tropical forage diets in the 60–70% DMD range without the GP sampling method.

Calibrations

Calibration statistics for the full sample set (ALL), for the GP, VIVO, and FR-FEED sets and ALLlessGP, ALLlessVIVO and ALLlessFF are presented in Table 3. The scatter plot of predicted DMD versus reference DMD for the full calibration ALL.EQA is shown in Figure 2. *SEC* was lowest for the FR-FEED sample set (1.09) but this was associated with the narrow range in DMD values and the lesser diversity of sample types (sites, pasture species, seasons and years) compared with the other sample sets. There was a tendency for *SEC* to increase with the size of the sample set (Table 3) and this also would probably be associated with greater diversity among calibration samples.

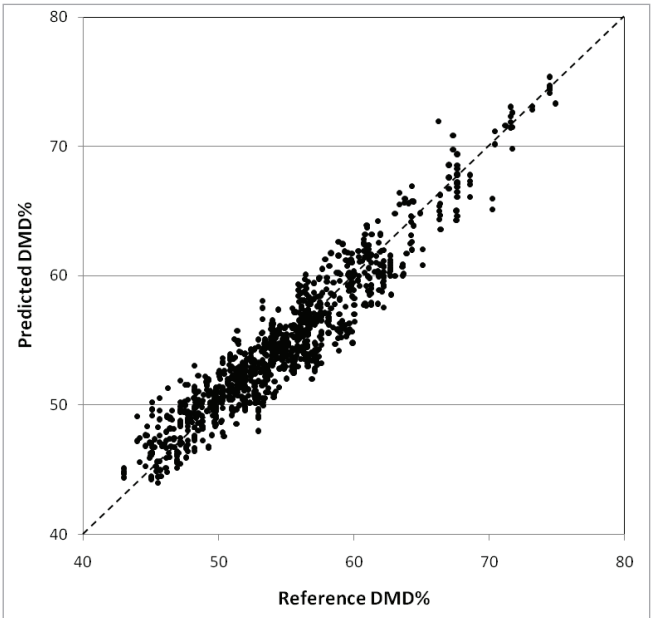

Figure 2. Scatter plot of predicted dry matter digestibility (DMD) vs reference DMD for the main calibration ALL.EQA. $n = 1044$, $r^2 = 0.90$.

The results of the present study compare favourably with calibration statistics reported in a recent review by Dixon and Coates,[22] especially considering the wide diversity of forage diets included in our calibration sets (264 different diets overall and 116, 71 and 77 different diets in the GP, VIVO and FR-FEED sample sets, respectively). Of the 18 cattle calibrations presented in the above review, but not including the calibration of Coates,[26] sample numbers exceeded 200 in only two of the calibration sets and were < 100 in 11 calibrations. *SEC* values (per cent digestibility) ranged from 1.6 to 6.5 with mean and median values of 3.0 and 2.8, respectively, compared with *SEC* in the present study of 1.87% digestibility for the entire sample set and less for most of the sub-sets.

In developing F.NIRS calibrations for predicting diet attributes of cattle, a major constraint to predictive accuracy is the accuracy of reference values. Accurate diet reference values depend not only on accurate laboratory analysis but also on measures taken to avoid or minimise miss-match error. In regard to the latter issue, it is important to appreciate that the composition of a single, small sample of faeces from cattle consuming forage diets is determined by the total feed ingested over two to four days.[27] Even when cattle are fed forages in pens, there is likely to be some between-day variation in the composition of forage offered and forage ingested leading to some degree of miss-match error. The frequency and magnitude of miss-match errors are likely to be greater with diet–faecal pairs from grazed pasture than from pen fed cattle, due to differences between the composition of extrusa samples collected from OF cattle and the diet of the resident cattle. Correcting calibration reference values in the GP sample set for the effect of saliva on *in vitro* dry matter disappearance and for miss-match error resulted in an improvement in calibration *SEC* and r^2 (*SEC* of 1.53 vs 1.72, and r^2 of 0.91 vs 0.87 for corrected and uncorrected values. respectively). While these differences were not large, the concern is that errors in calibration reference values would carry over into prediction errors.

Calibration statistics for the FR-FEED set were excellent with a low *SEC* of 1.09. However, a major disadvantage of the FR-FEED sampling method was the absence of diets with high DMD. This was primarily a consequence of the mechanical harvesting of pasture. Grazing cattle select for the higher quality components such as leaf rather than stem and green rather than dry material. Mechanical harvesting of pasture resulted in diets of much lower quality and much lower leaf to stem ratios than would be selected by grazing cattle. This was clearly evidenced by samples with reference DMD below 50% accounting for 45% of samples in the FR-FEED calibration set compared with such diets accounting for only 4% of GP samples (Figure 1). Apart from a diet of just under 60% DMD made by mixing a high proportion of high quality lucerne hay with grass hay at Swans Lagoon, the highest reference values in the FR-FEED set were for *Urochloa mosambicencis* (59%) and *Cenchrus ciliaris* (58%) at Lansdown early in the wet season where regrowth after mowing was young and leafy. There were no samples in the FR-FEED set with DMD in the 60–75% range, whereas diets within this range were common in the GP set and accounted for approximately 25% of the samples (Figure 1). The absence of tropical forage diets with high DMD also occurred in the VIVO set where non-tropicals (C_3 grass hays, lucerne hay, and C_3 grass/lucerne mixtures) accounted for all diets with DMD > 60%.

Validations

Statistics for the various validation exercises, using independent validation sets that represented the different sampling methods and the different sampling sites, are presented in Table 4. The non-independent prediction statistics for the same sampling method sets and site sub-sets obtained by using the main calibration ALL.EQA are shown for comparison.

While the table of validation statistics includes *SEP*, bias, r^2, slope and ratio of prediction to deviation (*RPD*) values, we considered *SEP* to be the most informative statistic considering the structure of the various validation sets. The *RPD*, r^2 and, to a lesser extent slope, are strongly influenced by the range in reference values and, although the range in DMD of the full sample set was 37 percentage units, the majority of the validation sets in Table 4 had ranges of less than half that value. *RPD*, as defined by Williams,[28] is calculated by dividing the standard deviation (*SD*) of the reference values of the validation set by the *SEP*. Given the limited range and therefore low *SD* of reference values for many of the validation sets, we considered the *SD* of the entire sample set (5.79) as representing the *SD* of the open population for which the calibration was developed to be more appropriate for calculating *RPD*. Despite this method of calculation, the *RPD* values in Table 4 did not rate highly according to Williams' classification[28] with 14 of the 25 independent validation predictions classified as very poor (≤ 2.3), eight as poor

Table 4. Validation statistics to determine the effect of sampling method, the effect of site within sampling method (for grazed pasture and fresh feed pen trials), and the efficacy of the main calibration in relation to sampling method and sites.

Validation	*n*	Range in DMD%	*SEP*	Bias	r^2	Slope	*RPD*[a]	Ave *H*
1a. VIVO set by GP.EQA	296	44–75	3.75	−0.82	0.73	1.28	1.54	1.63
1b. VIVO set by FR-FEED.EQA	296	44–75	3.55	2.07	0.81	0.98	1.63	3.13
1c. VIVO set by ALLlessVIVO.EQA	296	44–75	2.52	0.22	0.86	1.05	2.30	1.36
1d. VIVO set by ALL.EQA	296	44–75	2.07	−0.01	0.90	0.97	2.80	—
2a. FR-FEED set by GP.EQA	288	38–60	3.98	−1.54	0.13	0.42	1.45	3.11
2b. FR-FEED set by VIVO.EQA	288	38–60	3.61	−2.33	0.59	0.62	1.60	1.68
2c. FR-FEED by ALLlessFR-FEED.EQA	288	38–60	2.62	−0.05	0.47	0.78	2.21	1.38
2d. FR-FEED set by ALL.EQA	288	38–60	1.72	−0.27	0.76	0.98	3.37	—
3a. GP set by VIVO.EQA	468	47–75	3.70	−0.92	0.53	0.80	1.56	4.17
3b. GP set by FR-FEED.EQA	468	47–75	3.89	1.85	0.68	0.69	1.49	5.10
3c. GP set by ALLlessGP.EQA	468	47–75	2.84	−0.48	0.70	0.91	2.04	3.16
3d. GP set by ALL.EQA	468	47–75	1.96	0.15	0.85	0.98	2.95	—
4a. KTH FF set by FFlessKTH.EQA	50	43–58	2.36	0.26	0.69	0.92	2.45	1.79
4b. KTH FF set by ALLlessKTH.EQA	50	43–58	1.70	−0.40	0.88	1.24	2.40	1.14
4c. KTH FF set by ALL.EQA	50	43–58	1.38	−0.12	0.91	1.14	4.19	—
5a. BCH FF set by FFlessBCH.EQA	31	46–49	3.58	−2.76	0.01	−0.04	1.62	2.06
5b. BCH FF set by ALLlessBCH.EQA	31	46–49	4.45	−3.52	0.09	−0.12	1.30	1.25
5c. BCH FF set by ALL.EQA	31	46–49	2.95	−1.50	0.06	−0.10	1.96	—
6a. SWL FF set by FFlessSWL.EQA	104	45–60	1.65	0.51	0.68	1.07	3.51	2.44
6b. SWL FF set by ALLlessSWL.EQA	104	45–60	1.84	0.32	0.57	0.97	3.15	1.44
6c. SWL FF set by ALL.EQA	104	45–60	1.51	−0.26	0.71	0.98	3.83	—
7a. LDN FF set by FFlessLDN.EQA	99	45–59	2.42	1.00	0.66	0.81	2.39	3.96
7b. LDN FF set by ALLlessLDN.EQA	99	45–59	1.99	−0.12	0.73	0.82	2.91	1.21
7c. LDN FF set by ALL.EQA	99	45–59	1.56	−0.03	0.81	0.95	3.71	—
8a. CDG GP set by GPlessCDG.EQA	72	48–61	2.41	−0.50	0.71	0.25	2.40	1.68
8b. CDG GP set by ALLlessCDG.EQA	72	48–61	2.77	0.80	0.60	1.10	2.09	1.47
8c. CDG GP set by ALL.EQA	72	48–61	2.16	0.34	0.76	1.17	2.68	—
9a. SPR GP set by GPlessSPR.EQA	69	48–62	2.19	1.54	0.84	0.82	2.64	1.85
9b. SPR GP set by ALLlessSPR.EQA	69	48–62	1.89	0.30	0.84	0.72	3.06	1.29
9c. SPR GP set by ALL.EQA	69	48–62	1.45	−0.12	0.85	0.87	3.99	—
10a. BP GP set by GPlessBP.EQA	55	47–68	3.25	−1.10	0.86	1.26	1.78	3.45
10b. BP GP set by ALLlessBP.EQA	55	47–68	2.98	0.04	0.84	0.94	1.94	2.07
10c. BP GP set by ALL.EQA	55	47–68	2.45	0.43	0.89	1.02	2.36	—
11a. LDN GP set by GPlessLDN.EQA	250	51–70	2.85	−0.23	0.54	0.79	2.03	1.66
11b. LDN GP set by ALLlessLDNGP.EQA	250	51–70	2.46	−0.65	0.67	0.88	2.35	1.74
11c. LDN GP set by ALL.EQA	250	51–70	1.94	0.07	0.77	0.96	2.98	—

[a]*RPD* calculated using the *SD* of the full sample set.
Refer to previous tables and text for explanations of abbreviations.

(> 2.3 to ⩽ 3.0), only three as fair (>3.0 to ⩽4.9) and none as good (> 4.9). We consider these classifications provide an inappropriate evaluation of F.NIRS calibrations for predicting DMD and we favour an evaluation based simply on the *SEP* as an indicator of predictive accuracy. This requires a decision to be made regarding the prediction error that can be tolerated relative to the purpose for which the predictions are to be used together with the probability of a prediction being within the designated

error. For example, an *SEP* of 2.0 DMD percentage units would indicate prediction error to be ⩽2 for ~67% of samples analysed and ⩽4 for ~95% of samples analysed.

Validation *SEP* values for the independent tests presented in Table 4 varied between 1.70 and 4.45 and compared favourably with those for cattle presented in the Dixon and Coates review[22] where *SEP* for DMD% or organic matter digestibility (%) averaged 6.2 (range 2.0–11.5) and where *SEP* for eight of the 10 studies in the above review[22] were ⩾4.7. Despite what we considered acceptable *SEP* values for 16 of the 19 independent validation tests in Table 4 (*SEP* ⩽ 3), there were some disturbing features which we discuss below, especially when the validation statistics were compared with calibration *SEC* and *SECV* values (Table 3). This highlights the limitations of the *SECV* statistic which often gives a favourable but unrealistic assessment of equation performance.[29]

Effect of sampling method

It was apparent that calibrations developed on samples from a particular sampling method were not reliably good predictors on samples from the other two sampling methods so that *SEP* values for all six combinations ranged from 3.55 to 3.98 (Table 4, rows 1a, 1b, 2a, 2b, 3a and 3b). The relatively poor predictive accuracy of the FR-FEED equation was not surprising considering the narrow DMD reference range with no values >60%, and the limited sample diversity in relation to sites, pasture species/mixtures and years. However, predictive accuracy of the calibrations developed on the VIVO and GP sample sets appeared to be no better when applied to samples from the other sampling methods. This observation suggests that it is inappropriate to rely entirely on calibrations from pen fed cattle and, in particular, from cattle fed hay diets. The average global *H* values of the GP sample set relative to the VIVO and FR-FEED sample sets, separately or combined, were all >3 (Table 4) and therefore more than the recommended maximum for predictive reliability.[30] The high average *H* values indicated appreciable spectral differences between samples obtained from grazing cattle and pen fed cattle even though many of the FR-FEED diets were freshly harvested green pasture. We therefore suggest that at least a proportion, and preferably a substantial proportion, of calibration samples should be generated from grazed pastures if the equation is to be applied to grazing cattle. Nevertheless, we also consider that diet–faecal pairs from penned cattle fed hay or fresh feed are likely to be of benefit in calibration sets to increase diet diversity, especially because OF cattle and suitable infrastructure for sampling grazed pasture are not available at many locations and also because diet reference values can usually be measured more accurately in pen-fed cattle. Our results showed a substantial improvement in *SEP* when predictions on samples from one sampling method were derived from the calibration developed from the sample sets of the other two methods combined (Table 4, rows 1c, 2c and 3c) and a further substantial improvement in predictive accuracy when *SEP* for a sampling method set was based on predictions using the combined equation ALL.EQA (Table 4, rows 1d, 2d and 3d).

Effect of site

When calibrations were developed on sub-sets from either the GP or FR-FEED sampling methods, validation tests indicated that when a calibration with acceptably low *SEC* and *SECV* was used to predict samples from a site not included in the calibration set, predictive accuracy was often disappointing. Despite the average *H* of the validation sets being generally <3, this occurred for all sites except the FR-FEED sub-set from Swans Lagoon where the *SEP* was a very acceptable 1.65 (Table 4, row 6a), and to a lesser extent for the GP sub-set from Springmount where the *SEP* was 2.19 (Table 4, row 9a). When the number and diversity of samples in the calibration were increased by including samples from all three sampling methods, less the relevant site validation sub-set, there was a moderate to substantial improvement in *SEP* values for five of eight cases (Table 4, validations 4b, 7b, 9b, 10b and 11b) and a deterioration in the other three cases (Table 4, validations 5b, 6b, and 8b). Overall, however, there was an encouraging degree of apparent robustness inherent in the combined calibration sets with *SEP* values <2 for four of the eight independent validations (Table 4, validations 4b, 6b, 7b and 9b). Exceptions were the Brunchilly FR-FEED sample sub-set (5b) and the GP sample sub-sets from Cardigan (8b), Brian Pastures (10b) and Lansdown (11b). We suggest that serious miss-match errors in reference values may have contributed to the higher *SEP* values for some of these sites. The higher than desirable *SEP* for the Lansdown GP sub-set may have been partly due to this sub-set representing about one quarter of the entire sample set. Most of the high residuals (>3 percentage units) for the Lansdown GP sub-set were associated with sown grass/stylo diets with high legume content and these were not sufficiently well represented in the remainder of the sample set. There was a further improvement in the relationship between predicted and reference DMD when samples from the validation sub-sets were included in the calibration but the higher *SEP* values for Brunchilly, Cardigan and Brian Pastures compared with the other sites reinforced the possibility of miss-match errors being a problem in these sample sub-sets.

Validations using Grp1, Grp2 and Grp3

Validation statistics for the sample groups selected by dividing the full calibration set into three groups of similar size are presented in Table 5. *SEP* values averaged 2.37 which was 27% higher than the *SEC* of the calibration developed on the entire sample set (ALL.EQA). We consider that this *SEP* gives a conservative but realistic measure of likely performance of the current calibration when applied to samples derived from diets, pasture systems and regions similar to those from which the current calibration samples were derived.

The frequency and importance of miss-match errors

It has often been stated that NIR spectroscopy predictions can only be as accurate as the accuracy of the laboratory analysis used to determine reference values for calibration. While this is not necessarily true if laboratory analysis errors are random and not biased,[31] it is still a useful principle because random

Table 5. Validation statistics for the three sub-sets (Grp 1, Grp 2 and Grp 3), each approximately one third of the entire calibration set and each predicted with the calibration equation developed on the other two sub-sets combined.

Validation set	*n*	Range in reference DMD%	*SEP*	Bias	r^2	Slope	*RPD*[a]	Ave *H*
Grp 1	358	43.0–71.4	2.37	–0.31	0.80	0.98	2.44	1.06
Grp 2	329	44.2–68.6	2.19	–0.07	0.84	0.95	2.64	1.03
Grp 3	364	44.7–74.9	2.55	0.44	0.85	0.99	2.27	1.22
Mean	—	—	2.37	—	0.83	—	2.45	—

[a]*RPD* calculated using the *SD* of the full sample set
See text for allocation of samples to the different sub-sets.

errors in reference values result in poorer calibration and validation statistics. Importantly, however, accuracy of laboratory analysis will be of little avail if there is any miss-matching in composition between the sample on which the laboratory analysis is conducted and the sample responsible, directly or indirectly, for the NIR data which, in this study, is the diet of the cattle from which faeces are collected. In most NIR applications where the laboratory analysis and NIR analysis are conducted on the same sample, miss-match errors should be minor, provided appropriate sub-sampling procedures are used. However, in the development of F.NIRS calibration equations for measuring dietary attributes (as distinct from faecal attributes) the reference values are determined on samples collected to represent the diet while the NIR spectra are measured on the faeces. When the diet is either hay or freshly harvested forage fed to penned cattle, or when diet samples are obtained from OF cattle grazing pasture, it is not possible to avoid some miss-match error and such errors may be large.

Given the risk of miss-match errors, it is clearly important to try and detect those samples where errors are of such a magnitude that they have serious adverse effects on calibration statistics and equation performance. In some cases, faecal spectra associated with large miss-match errors will be identified as outliers and eliminated during calibration. In other cases, however, miss-match errors of a sufficient number may lead to bias or just to poor calibration statistics. For example, if the reference values are under-estimated (or over-estimated) for sufficient samples in the calibration set, this is likely to lead to bias in the prediction of samples of a certain type. Observations during field sampling to develop the GP data set suggested that miss-match errors did occur. For example, at the Cardigan site, when extrusa was being collected from OF steers grazing a particular native pasture paddock, it was observed that they primarily grazed large, stemmy tussocks of desert bluegrass (*Bothriochloa ewartiana*) whereas it was obvious that the resident steers exerted heavy grazing pressure along a drainage line which the OF steers avoided. Due to this observed grazing behaviour, we anticipated on a number of sampling occasions that the extrusa collected was likely to be of lower DMD than the diet of the resident steers. When these resident faecal samples were predicted using calibrations independent of the Cardigan site, predicted values were indeed consistently higher than the reference values measured on the extrusa, thus supporting the hypothesis of likely miss-match error. This phenomenon also appeared to occur for diet–faecal pairs from grazed mature native pasture at the Springmount and Hillgrove sites, suggesting that the resident cattle were more selective than the OF steers. Even when these samples were represented in the calibration equation, the predicted values were higher than the reference values and, therefore, consistent with the hypothesis for miss-match error. Where the evidence for serious miss-match error was sufficiently strong, it was deemed prudent to exclude such samples from the calibration set. There were similar occasions where serious miss-match errors likely occurred in FR-FEED diet–faecal pairs when native pasture was fed to cattle at Brunchilly and Katherine. In these cases, it appeared that the forage consumed by the cattle was of higher DMD than the forage samples analysed, most likely as a result of selection by the penned cattle and/or due to errors in sub-sampling the forage for laboratory analysis. Where there was convincing evidence for substantial miss-match error, the samples were excluded from the calibration set.

Despite the elimination from the calibration set of some diet–faecal pairs due to probable miss-match error, it is likely that many samples retained in the calibration set were also affected by miss-match error. Therefore, we suggest that the derived calibrations are likely to be somewhat better than the calibration statistics indicate, since such statistics attribute zero error to the reference values. Moreover, we also speculate that miss-match error probably contributed to many of the larger residuals (laboratory minus predicted) that were observed. For example, although a number of samples from Brunchilly were discarded from the calibration set on the evidence of miss-match error, the plots of predicted DMD against reference DMD strongly suggested that some of the samples retained were also suspect (Figure 3). The plot shows that when Brunchilly samples were included in the FR-FEED calibration set (31 samples in a total of 288), the MPLS regression process was able to accommodate the spectral differences to obtain an acceptable fit between reference and predicted values for the Brunchilly samples [Figure 3(a)]. However, when the Brunchilly samples were excluded from the calibration set in FFlessBCH.EQA, there were many large residuals when the equation was

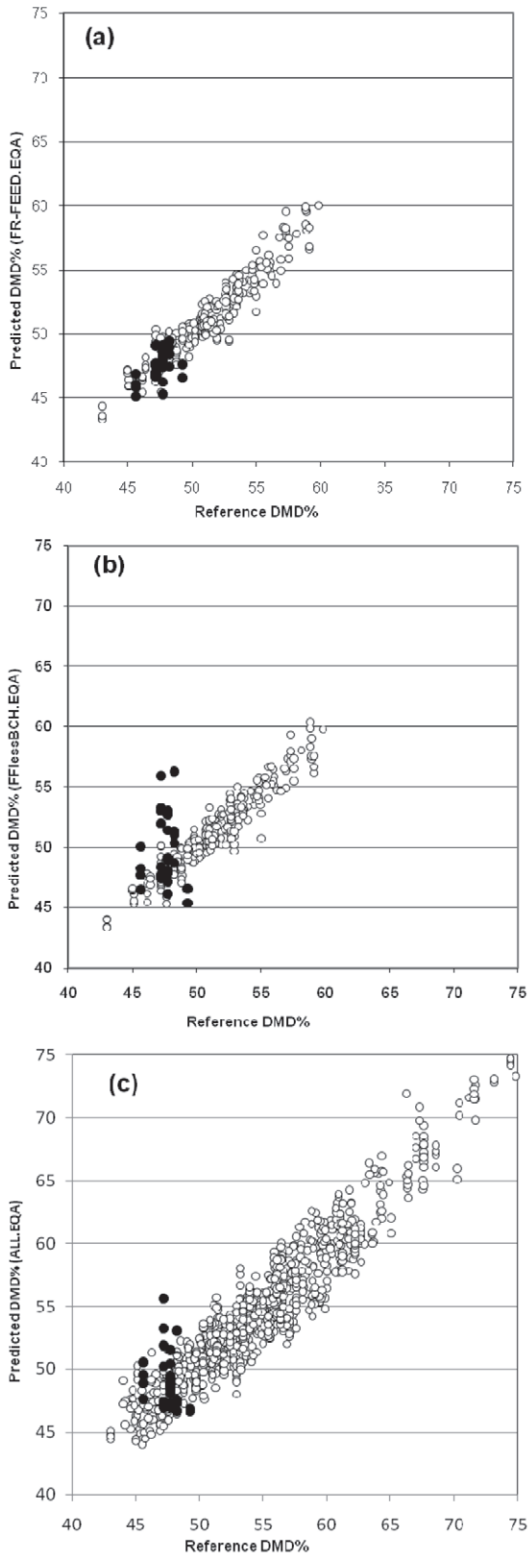

Figure 3. Predicted DMD% plotted against reference DMD%. (a) FR-FEED sample set predicted with equation FR-FEED.EQA (with Brunchilly samples included in calibration). (b) FR-FEED sample set predicted with equation FFlessBCH.EQA (Brunchilly samples not included in calibration). (c) Entire sample set predicted with ALL.EQA (with Brunchilly samples included in calibration). (●) Brunchilly samples; (○) other samples.

applied to predict the Brunchilly samples [Figure 3(b)]. In addition, even when the Brunchilly samples were included in the combined calibration set in equation ALL.EQA (31 samples in a total of 1052), the differences between reference and predicted DMD for the Brunchilly samples were similar in magnitude to those from the predictions using FFlessBCH.EQA [Figure 3(c)]. It seems reasonable to suspect that the large residuals were probably associated with serious miss-match error and that many of the larger differences between reference and predicted values in ALL.EQA were also the consequence of miss-match. It is noteworthy that the two Brunchilly samples with the largest prediction errors [Figure 3(c)] were eliminated as outliers during calibration.

It is not possible to determine the extent of miss-match error but if, as we suspect, miss-match errors are a significant problem, it follows that that the predictive accuracy of the various calibration equations used in this study was likely to be somewhat better than first indicated by the calibration statistics in Table 3 and, more importantly, by the validation statistics shown in Tables 4 and 5. Consequently, we strongly recommend that appropriate sampling procedures and quality control measures, based on an awareness of the high risk of miss-match errors, and the need to minimise such errors, should receive serious consideration in F.NIRS involving diet–faecal pairs if predictive accuracy is to be achieved and if validation tests are to be meaningful.

Overall assessment of predictive performance

Unlike most applications of NIR spectroscopy in agriculture, it is time-consuming and costly to monitor the performance of F.NIRS calibrations that predict dietary attributes. This is because reference values for monitoring predictive performance can only be determined on the diet which necessitates additional well-managed experiments to be conducted. Nevertheless, it is well accepted that confidence in NIRS predictive reliability requires ongoing monitoring and updating of calibration equations. Despite the encouraging characteristics of the main calibration described in this paper, especially when compared with those from previous studies,[22] we believe there is potential to improve robustness by substantially expanding the number and diversity of diets in the calibration set to cope with the range of conditions and pasture systems that exist across northern Australia.

Conclusions

The present study provides evidence that F.NIRS offers a reliable technology to estimate DMD of forage diets selected by grazing cattle in tropical and sub-tropical rangelands of northern Australia. Predictive accuracy is likely to be acceptable for many research and grazing management applications for those pasture systems well represented in the calibration data set.

Although the entire calibration set of 1052 samples was derived from 264 different diets, some northern pasture

systems are not well represented, or not represented at all. Thus, the current calibration may not be sufficiently robust to accurately predict all samples from cattle grazing rangelands in northern Australia.

Evidence from the current study and associated research[24,25] indicates special attention is required with respect to sampling methods in the construction of appropriate calibration data and with respect to determining accurate DMD reference values for pairing with faecal spectra.

For calibration equations to be applied to grazing cattle, predictive accuracy requires that a substantial proportion of calibration faecal spectra be obtained from grazing rather than pen fed cattle. Faecal spectra from cattle fed processed forages such as hay can be usefully included in calibration sets but calibrations developed entirely on hay diets apparently do not function well when applied to faecal spectra of grazing cattle (and *vice versa*). Similarly, although the feeding of mechanically harvested forage fed straight from the paddock provides a useful sampling method to increase diet diversity in the calibration set, this sampling method has serious limitations and should not be used solely, or even in conjunction with forage hays, to develop calibrations for use with grazing cattle.

As with all NIR applications predictive accuracy requires accurate reference values for pairing with calibration spectra. This presents special difficulties for F.NIRS prediction of DMD with respect to both the analytical procedure and to potential miss-match error. Where OF animals are used to provide diet samples for *in vitro* analysis, due consideration should be given to differences between extrusa and forage samples in dry matter or organic matter loss during *in vitro* incubation and the need to adjust extrusa values accordingly as recommended by Coates and Mayer.[24] In addition it is necessary to use the best possible mathematical relationships available to convert *in vitro* dry matter loss into estimated *in vivo* DMD to obtain accurate and consistent reference values.

Miss-match errors in diet–faecal pairs pose a serious threat to equation statistics, equation performance and meaningful validation statistics. For our calibration set the risk was lowest for the VIVO sampling method, highest for the GP sampling method due to extrusa collected from OF cattle not being reliably representative of the diet of resident cattle and intermediate for the FR-FEED sampling method. Experimental procedures should be designed to minimise miss-match errors. We suggest our efforts to minimise reference errors likely made a useful contribution to our calibration and validation statistics comparing favourably with other published reports.

Acknowledgements

We wish to acknowledge the funding support provided by MLA and the assistance provided by various technical staff within CSIRO, the Queensland Department of Primary Industries (now the Department of Employment, Economic Development and Innovation), the N.T. Department of Business, Industry & Resource Development, and access to information from other research programs conducted by research staff at the University of Queensland and the Queensland Department of Primary Industries.

References

1. D.J. Minson, *Forage in Ruminant Nutrition*. Academic Press, New York, USA (1990).
2. J.M.A. Tilley and R.A. Terry, "A two-stage technique for the *in vitro* digestion of forage crops", *J. Br. Grassland Soc.* **18,** 104 (1963). doi: 10.1111/j.1365-2494.1963.tb00335.x
3. D.I.H. Jones and M.V. Hayward, "The effect of pepsin pre-treatment of herbage on the prediction of dry matter digestibility from solubility in fungal cellulase solutions", *J. Sci. Food Agric.* **26,** 711 (1975). doi: 10.1002/jsfa.2740260518
4. M.N. McLeod and D.J. Minson, "The accuracy of the pepsin–cellulase technique for estimating the dry matter digestibility *in vivo* of grasses and legumes", *Anim. Feed Sci. Technol.* **3,** 277 (1978). doi: 10.1016/0377-8401(78)90002-0
5. E.R. Orskov, "The *in situ* technique for the estimation of forage degradability in ruminants", in *Forage Evaluation in Ruminant Nutrition*, Ed by D.I. Givens, E. Owens, R.F.E. Axford and H.M. Omed. CAB International, Wallingford, UK (2000).
6. S. Lopez, "*In vitro* and *in situ* techniques for estimating digestibility", in *Quantitative Aspects of Ruminant Digestion and Metabolism*, 2nd Edn. Ed by J. Dijkstra, J.M. Forbes and J. France. CAB International, Wallingford, UK (2005).
7. D.J. Minson, "Effect of chemical composition on feed digestibility and metabolizable energy", *Nutr. Abstract Rev. B* **52,** 591 (1982).
8. M.D. Stern, A. Bach and S. Calsamiglia, "Alternative techniques for measuring nutrient digestion in ruminants", *J. Anim. Sci.* **75,** 2256 (1997).
9. D.E. Beever and F.L. Mould, "Forage evaluation for efficient ruminant livestock production", in *Forage Evaluation in Ruminant Nutrition*. Ed by D.I. Givens, E. Owen, R.F.E. Axford and H.M. Omed. CAB International, Wallingford, UK (2000). doi: 10.1079/9780851993447.0015
10. G.M. Van Dyne and D.T. Torrel, "Development and use of the oesophageal fistula: A review", *J. Range Manage.* **17,** 7 (1964). doi: 10.2307/3895540
11. W.R. McManus, "Oesophageal fistulation techniques as an aid to diet evaluation of the grazing ruminant" in *Forage Evaluation: Concepts and Techniques*, Ed by J.L. Wheeler and R.D. Mochrie. CSIRO Australia and the American Forage and Grassland Council, Melbourne, Australia, p. 249 (1981).
12. D.B. Coates, P. Schachenmann and R.J. Jones, "Reliability of extrusa samples collected from steers

fistulated at the oesophagus to estimate the diet of resident animals in grazing experiments", *Aust. J. Exp. Agric.* **27,** 739 (1987). doi: 10.1071/EA9870739

13. R.J. Jones and C.E. Lascano, "Oesophageal fistulated cattle can give unreliable estimates of the proportion of legume in the diets of resident animals grazing tropical pastures", *Grass Forage Sci.* **47,** 128 (1992). doi: 10.1111/j.1365-2494.1992.tb02255.x

14. R.J. Clements, R.M. Jones, L.R. Valdes and G.A. Bunch, "Selection of *Chamaecrista rotundifolia* by cattle", *Trop. Grasslands* **30,** 389 (1996).

15. G.D. Barber, D.I. Givens, M.S. Kridis, N.W. Offer and I. Murray, "Prediction of the organic matter digestibility of grass silage", *Anim. Feed Sci. Technol.* **28,** 115 (1990). doi: 10.1016/0377-8401(90)90074-I

16. D.I. Givens and E.R. Deaville, "The current and future role of near infrared reflectance spectroscopy in animal nutrition: a review", *Aust. J. Agric. Res.* **50,** 1131 (1999). doi: 10.1071/AR98014

17. C.A. Roberts, J. Stuth and P. Flinn, "Analysis of forages and feedstuffs", in *Near-Infrared Spectroscopy in Agriculture*, Agronomy Monograph 44. American Society of Agronomy, Madison, Winsconsin, USA. p. 231 (2004).

18. S. Andres, I. Murray, A. Calleja and F.J. Giraldez, "Review. Nutritive evaluation of forages by near infrared reflectance spectroscopy", *J. Near Infrared Spectrosc.* **13,** 301 (2005). doi: 10.1255/jnirs.561

19. J.W. Stuth, E.D. Kapes and R.K. Lyons, "Use of near infrared spectroscopy to assess nutritional status of cattle diets on rangeland", in *Proceedings of the XVI International Grassland Congress, Nice, France*. The French Grassland Society, Versailles, France, p. 889 (1989).

20. R.K. Lyons and J.W. Stuth, "Fecal NIRS equations for predicting diet quality of free-ranging cattle", *J. Range Manage.* **45,** 238 (1992). doi: 10.2307/4002970

21. E.R. Leite and J.W. Stuth, "Faecal NIRS equations to assess diet quality of free-ranging goats", *Small Rum. Res.* **15,** 223 (1995). doi: 10:1016/0921-4488(94)00026-4

22. R.M. Dixon and D.B. Coates, "Review: Near infrared spectroscopy of faeces to evaluate the nutrition and physiology of herbivores", *J. Near Infrared Spectrosc.* **17,** 1 (2009). doi: 10.1255/jnirs.822

23. V. Decruyenaere, P. Lecomte, C. Demarquilly, J. Aufrere, P. Dardenne, D. Stilmant and A. Buldgen, "Evaluation of green forage intake and digestibility in ruminants using near infrared reflectance spectroscopy (NIRS): Developing a global calibration", *Anim. Feed Sci. Technol.* **148,** 138 (2009). doi: 10.1016/j.anifeedsci.2008.03.007

24. D.B. Coates and R.J. Mayer, "Differences between the *in vitro* digestibility of extrusa collected from oesophageal fistulated steers and the forage consumed", *Anim. Prod. Sci.* **49,** 563 (2009). doi: 10.1071/EA08285

25. D.B. Coates, " Improving the accuracy of digestibility estimates for cattle grazing tropical pastures when sampling with oesophageal fistulated cattle", *Anim. Prod. Sci.* **51,** 44 (2011). doi: 10.1071/AN10131

26. D.B. Coates, *Faecal NIRS—Technology for Improving Nutritional Management of Grazing Cattle*, Final Report of Project NAP3.121. Meat and Livestock Australia, Sydney, Australia (2004).

27. D.B. Coates, A.P.A. Van der Weide and J.D. Kerr, "Changes in faecal $\delta^{13}C$ in response to changing proportions of legume (C3) and grass (C4) in the diet of sheep and cattle", *J. Agric. Sci. Cambridge* **116,** 287 (1991).

28. P.C. Williams, "Implementation of near-infrared technology", in *Near-Infrared Technology in the Agricultural and Food Industries*, 2nd Edn, Ed by P. Williams and K. Norris. American Association of Cereal Chemists, Inc. St Paul, Minnesota, USA (2001).

29. D.B. Coates and R.M. Dixon, "Development of near infrared analysis of faeces to estimate non-grass proportions in diets selected by cattle grazing tropical pastures", *J. Near Infrared Spectrosc.* **16,** 471 (2008). doi: 10.1255/jnirs.815

30. J.S. Shenk and M.O. Westerhaus, *Analysis of Agricultural and Food Products by Near Infrared Reflectance Spectroscopy*. Infrasoft International, NIRSystems, Port Matilda, USA (1993).

31. D.B. Coates, "Is near infrared spectroscopy only as good as the laboratory reference values? An empirical approach", *Spectrosc. Eur.* **14(4),** 24 (2002).